Full wwPDB NMR Structure Validation Report ⓘ

Apr 21, 2024 – 01:09 AM EDT

PDB ID : 2L2B
Title : Structure of StnII-Y111N, a mutant of the sea anemone actinoporin Sticholysin II
Authors : Pardo-Cea, M.A.; Bruix, M.; Santoro, J.
Deposited on : 2010-08-16

The following versions of software and data (see references ⓘ) were used in the production of this report:

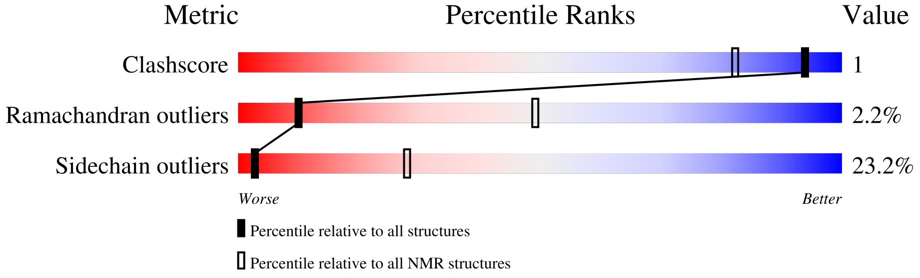
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	175	

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 12 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:2-A:175 (174)	0.66	12

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters and 4 single-model clusters were found.

Cluster number	**Models**
1	1, 2, 5, 8, 10, 12, 14, 15, 17, 19
2	3, 7, 11, 16
3	18, 20
Single-model clusters	4; 6; 9; 13

3 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 2689 atoms, of which 1332 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Sticholysin-2.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	175	2689	866	1332	230	255	6	0

There is a discrepancy between the modelled and reference sequences:

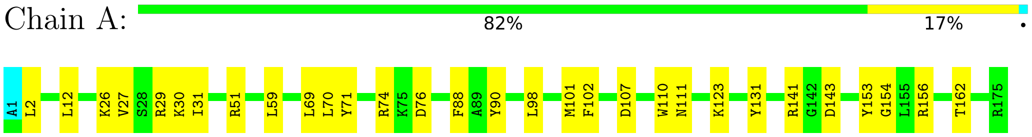
Chain	Residue	Modelled	Actual	Comment	Reference
A	111	ASN	TYR	engineered mutation	UNP P07845

4 Residue-property plots ⓘ

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: Sticholysin-2

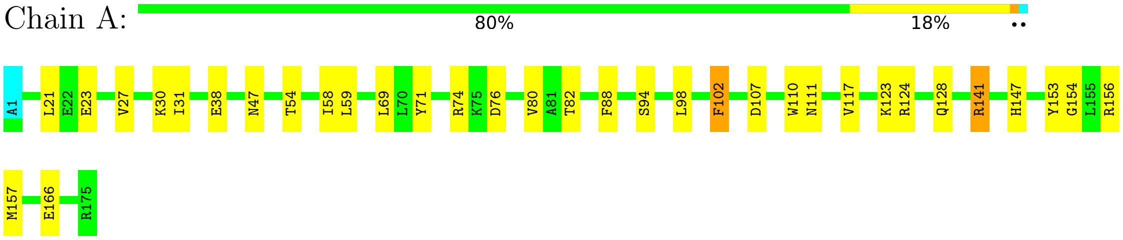


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

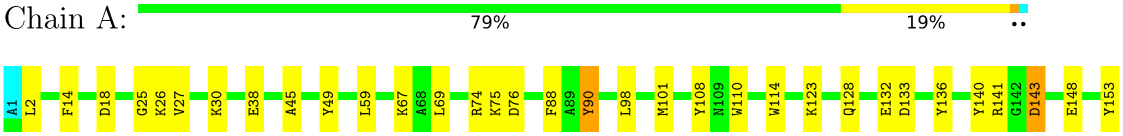
4.2.1 Score per residue for model 1

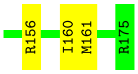
- Molecule 1: Sticholysin-2



4.2.2 Score per residue for model 2

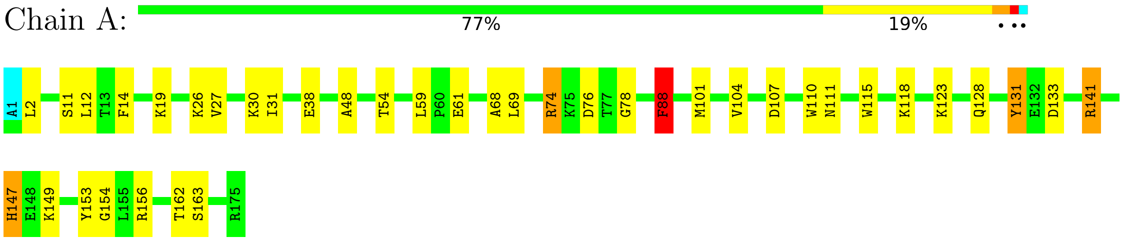
- Molecule 1: Sticholysin-2





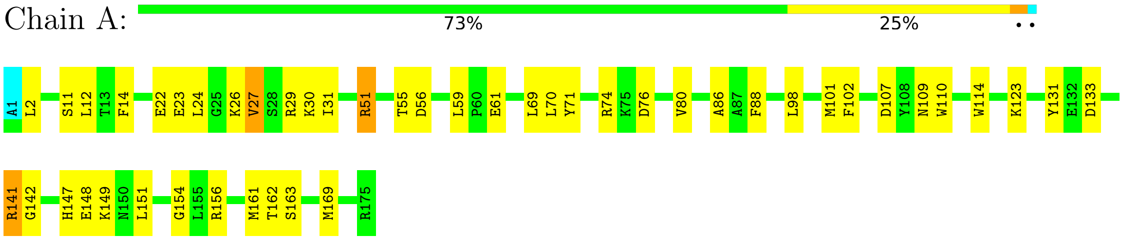
4.2.3 Score per residue for model 3

- Molecule 1: Sticholysin-2



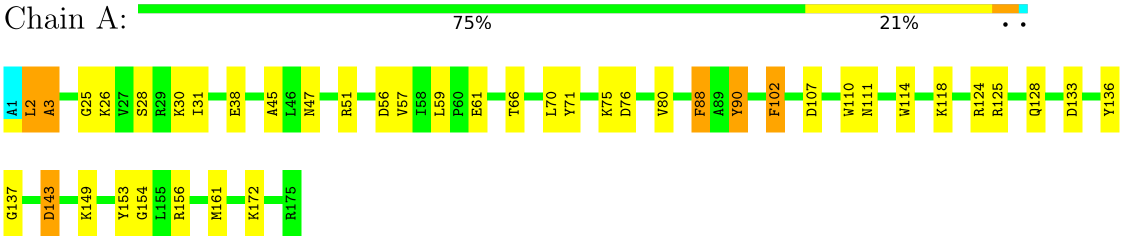
4.2.4 Score per residue for model 4

- Molecule 1: Sticholysin-2



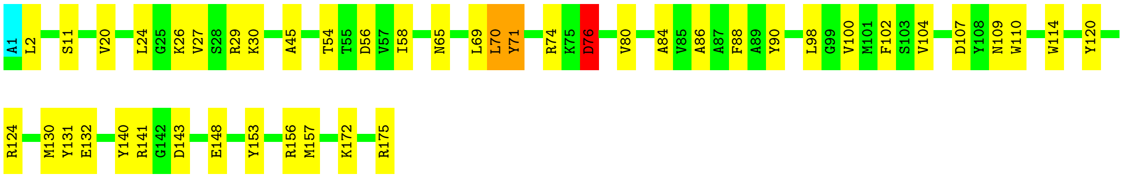
4.2.5 Score per residue for model 5

- Molecule 1: Sticholysin-2



4.2.6 Score per residue for model 6

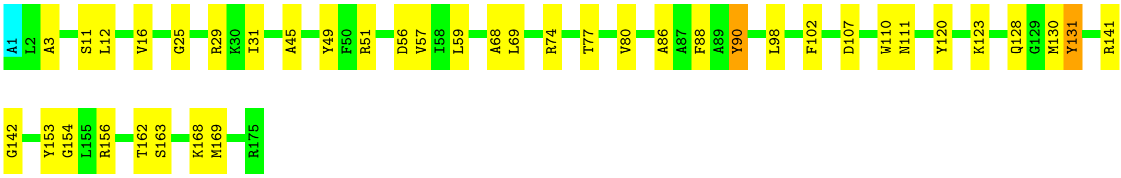
- Molecule 1: Sticholysin-2



4.2.7 Score per residue for model 7

- Molecule 1: Sticholysin-2

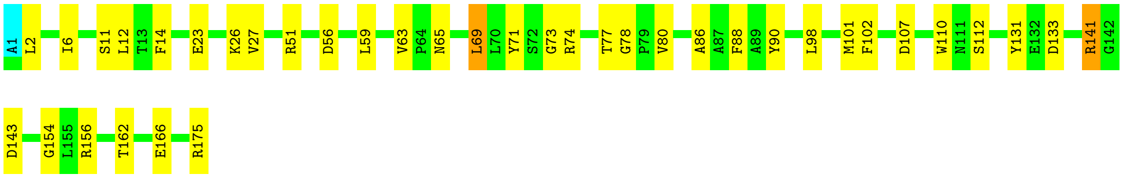
Chain A:



4.2.8 Score per residue for model 8

- Molecule 1: Sticholysin-2

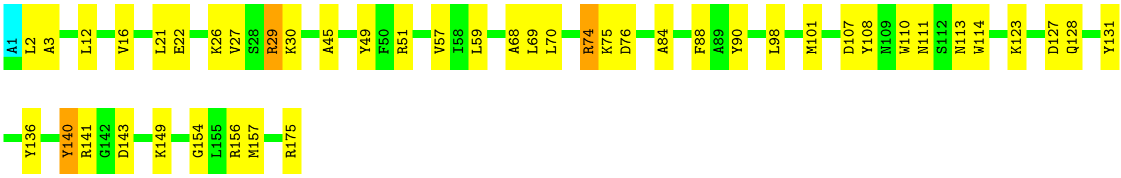
Chain A:



4.2.9 Score per residue for model 9

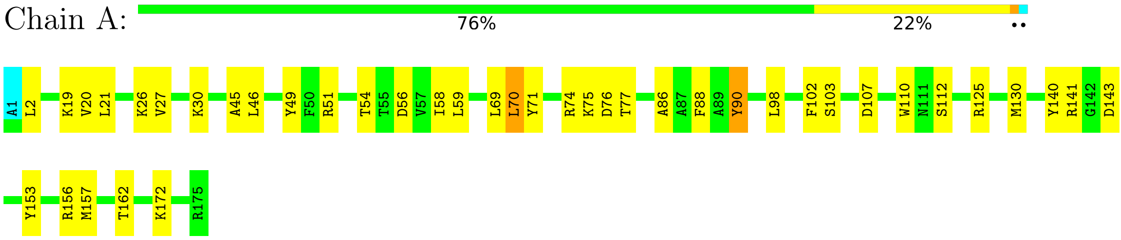
- Molecule 1: Sticholysin-2

Chain A:



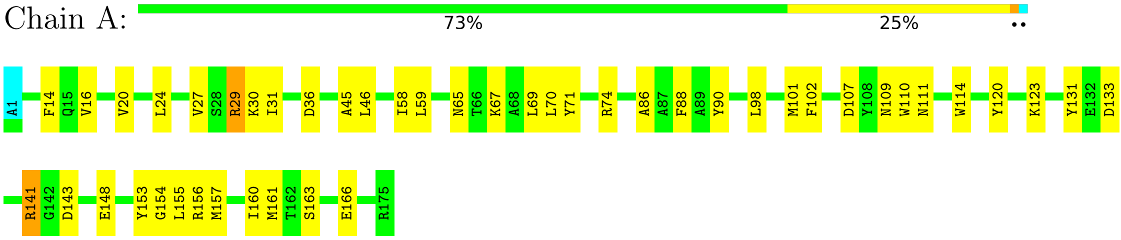
4.2.10 Score per residue for model 10

- Molecule 1: Sticholysin-2



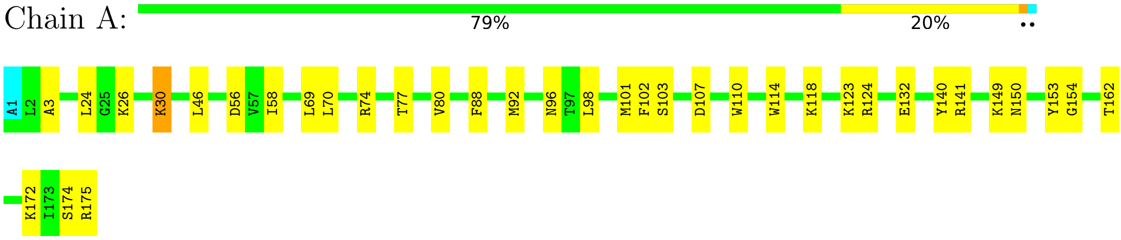
4.2.11 Score per residue for model 11

- Molecule 1: Sticholysin-2



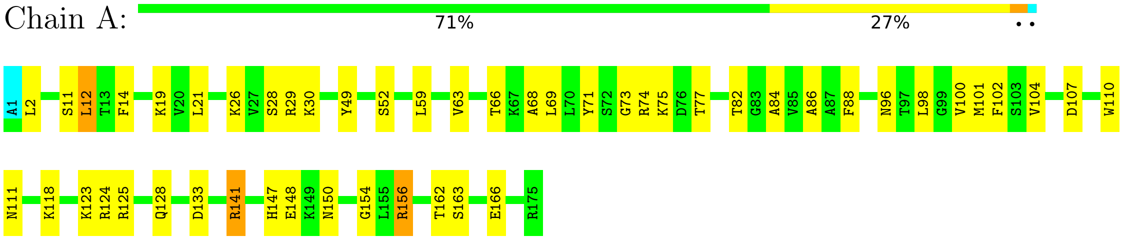
4.2.12 Score per residue for model 12 (medoid)

- Molecule 1: Sticholysin-2



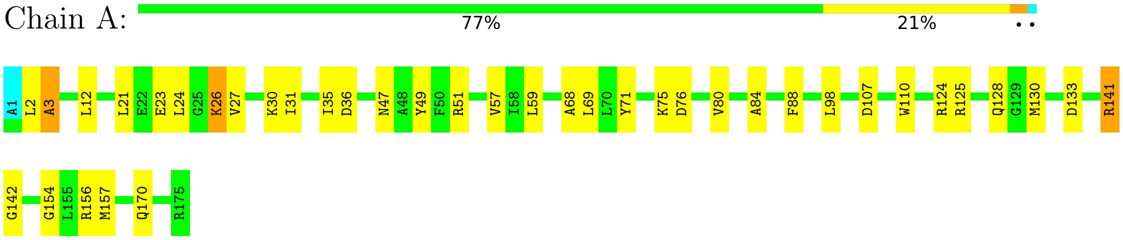
4.2.13 Score per residue for model 13

- Molecule 1: Sticholysin-2



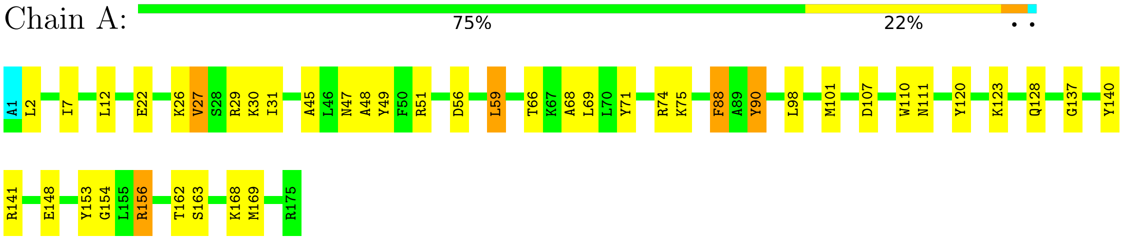
4.2.14 Score per residue for model 14

- Molecule 1: Sticholysin-2



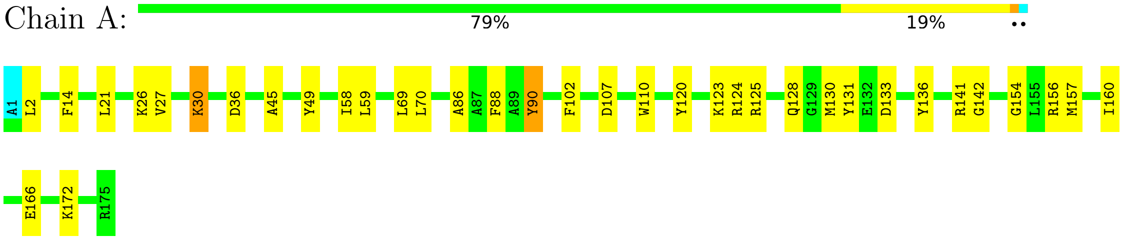
4.2.15 Score per residue for model 15

- Molecule 1: Sticholysin-2



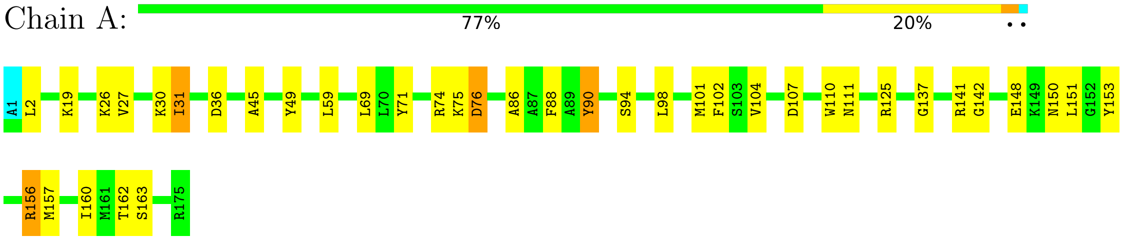
4.2.16 Score per residue for model 16

- Molecule 1: Sticholysin-2



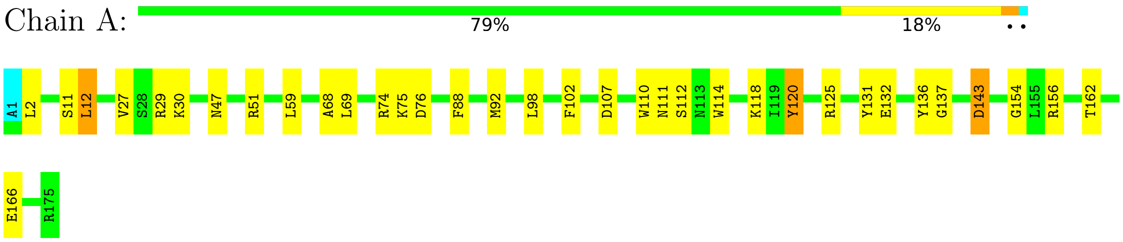
4.2.17 Score per residue for model 17

- Molecule 1: Sticholysin-2



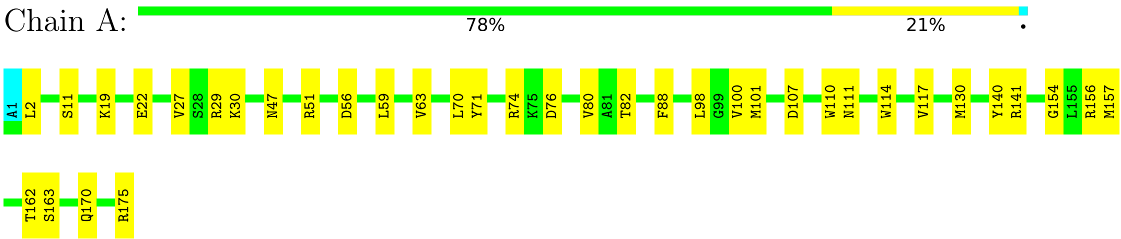
4.2.18 Score per residue for model 18

- Molecule 1: Sticholysin-2



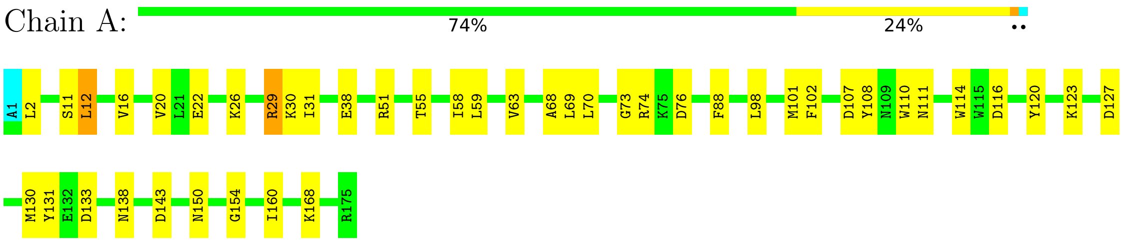
4.2.19 Score per residue for model 19

- Molecule 1: Sticholysin-2



4.2.20 Score per residue for model 20

- Molecule 1: Sticholysin-2



5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *molecular dynamics*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
Amber	refinement	

No chemical shift data was provided.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.76±0.00	0±0/1384 (0.0± 0.0%)	1.02±0.01	1±1/1871 (0.1± 0.1%)
All	All	0.76	0/27680 (0.0%)	1.02	22/37420 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.8±0.6
All	All	0	15

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	29	ARG	NE-CZ-NH1	7.61	124.11	120.30	9	8
1	A	29	ARG	CD-NE-CZ	6.66	132.92	123.60	9	1
1	A	156	ARG	CD-NE-CZ	5.50	131.30	123.60	13	1
1	A	141	ARG	NE-CZ-NH1	5.46	123.03	120.30	8	1
1	A	140	TYR	CB-CG-CD1	-5.27	117.84	121.00	9	1
1	A	136	TYR	CB-CG-CD2	-5.27	117.84	121.00	9	5
1	A	156	ARG	NE-CZ-NH1	5.21	122.91	120.30	17	2
1	A	51	ARG	NE-CZ-NH1	5.18	122.89	120.30	4	1
1	A	88	PHE	CB-CG-CD2	-5.10	117.23	120.80	5	2

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	71	TYR	Sidechain	8
1	A	131	TYR	Sidechain	4
1	A	120	TYR	Sidechain	2
1	A	74	ARG	Sidechain	1

6.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1352	1325	1324	3±1
All	All	27040	26500	26480	64

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:31:ILE:HD13	1:A:102:PHE:CD2	0.69	2.23	1	2
1:A:86:ALA:HB3	1:A:102:PHE:CZ	0.60	2.30	7	9
1:A:45:ALA:HB2	1:A:90:TYR:CZ	0.59	2.33	17	10
1:A:20:VAL:HG21	1:A:70:LEU:HD21	0.55	1.78	10	1
1:A:31:ILE:HD12	1:A:169:MET:SD	0.55	2.41	4	1
1:A:76:ASP:HB3	1:A:80:VAL:HG21	0.55	1.79	6	1
1:A:84:ALA:HB3	1:A:104:VAL:HB	0.55	1.79	13	2
1:A:3:ALA:CB	1:A:57:VAL:HG11	0.53	2.33	5	3
1:A:24:LEU:HD12	1:A:30:LYS:HE3	0.53	1.81	14	1
1:A:20:VAL:HG21	1:A:70:LEU:HD11	0.52	1.82	6	1
1:A:21:LEU:HD23	1:A:30:LYS:HD2	0.51	1.79	16	1
1:A:12:LEU:HD22	1:A:68:ALA:CB	0.50	2.37	18	8
1:A:16:VAL:O	1:A:20:VAL:HG23	0.50	2.07	20	2
1:A:3:ALA:HB3	1:A:57:VAL:HG11	0.48	1.85	5	2
1:A:73:GLY:HA3	1:A:102:PHE:CZ	0.48	2.44	13	2
1:A:3:ALA:CB	1:A:24:LEU:HD12	0.47	2.39	12	1
1:A:31:ILE:HG21	1:A:104:VAL:HG21	0.47	1.85	3	1
1:A:20:VAL:HG21	1:A:70:LEU:HD12	0.47	1.87	11	1
1:A:69:LEU:HD22	1:A:90:TYR:CZ	0.45	2.45	8	1
1:A:115:TRP:CZ2	1:A:147:HIS:CD2	0.44	3.05	3	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:31:ILE:CG2	1:A:104:VAL:HG21	0.44	2.43	3	2
1:A:74:ARG:C	1:A:84:ALA:HB2	0.43	2.34	9	1
1:A:6:ILE:HG21	1:A:63:VAL:HG22	0.43	1.91	8	1
1:A:75:LYS:HA	1:A:84:ALA:HB2	0.43	1.89	14	1
1:A:24:LEU:HD13	1:A:30:LYS:NZ	0.42	2.29	12	1
1:A:48:ALA:HB2	1:A:59:LEU:HD22	0.42	1.91	15	1
1:A:31:ILE:HD11	1:A:102:PHE:CZ	0.41	2.50	7	1
1:A:48:ALA:HB1	1:A:88:PHE:CD1	0.41	2.50	3	2
1:A:55:THR:HG22	1:A:73:GLY:HA2	0.41	1.92	20	1
1:A:3:ALA:HB3	1:A:24:LEU:HD12	0.40	1.93	12	1
1:A:100:VAL:HG12	1:A:117:VAL:HG13	0.40	1.92	19	1

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	173/175 (99%)	148±3 (86±2%)	21±3 (12±2%)	4±1 (2±1%)	10	49
All	All	3460/3500 (99%)	2970 (86%)	413 (12%)	77 (2%)	10	49

All 14 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	154	GLY	16
1	A	76	ASP	13
1	A	143	ASP	7
1	A	141	ARG	6
1	A	27	VAL	5
1	A	142	GLY	5
1	A	26	LYS	4
1	A	137	GLY	4
1	A	77	THR	4
1	A	12	LEU	4
1	A	25	GLY	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	3	ALA	3
1	A	78	GLY	2
1	A	166	GLU	1

6.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	138/138 (100%)	106±4 (77±3%)	32±4 (23±3%)	3 28
All	All	2760/2760 (100%)	2119 (77%)	641 (23%)	3 28

All 101 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	88	PHE	20
1	A	110	TRP	20
1	A	107	ASP	19
1	A	59	LEU	18
1	A	69	LEU	18
1	A	156	ARG	18
1	A	30	LYS	17
1	A	74	ARG	17
1	A	98	LEU	17
1	A	141	ARG	17
1	A	2	LEU	16
1	A	27	VAL	12
1	A	111	ASN	12
1	A	123	LYS	12
1	A	101	MET	12
1	A	26	LYS	12
1	A	153	TYR	11
1	A	162	THR	11
1	A	51	ARG	11
1	A	128	GLN	10
1	A	114	TRP	10
1	A	133	ASP	10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	157	MET	9
1	A	49	TYR	9
1	A	11	SER	9
1	A	56	ASP	9
1	A	70	LEU	9
1	A	80	VAL	8
1	A	75	LYS	8
1	A	131	TYR	8
1	A	163	SER	8
1	A	58	ILE	7
1	A	124	ARG	7
1	A	14	PHE	7
1	A	90	TYR	7
1	A	140	TYR	7
1	A	148	GLU	7
1	A	125	ARG	7
1	A	130	MET	7
1	A	47	ASN	6
1	A	120	TYR	6
1	A	21	LEU	5
1	A	38	GLU	5
1	A	71	TYR	5
1	A	102	PHE	5
1	A	166	GLU	5
1	A	143	ASP	5
1	A	160	ILE	5
1	A	19	LYS	5
1	A	118	LYS	5
1	A	149	LYS	5
1	A	22	GLU	5
1	A	172	LYS	5
1	A	29	ARG	5
1	A	175	ARG	5
1	A	31	ILE	5
1	A	23	GLU	4
1	A	54	THR	4
1	A	147	HIS	4
1	A	132	GLU	4
1	A	161	MET	4
1	A	36	ASP	4
1	A	150	ASN	4
1	A	82	THR	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	108	TYR	3
1	A	61	GLU	3
1	A	24	LEU	3
1	A	109	ASN	3
1	A	66	THR	3
1	A	65	ASN	3
1	A	168	LYS	3
1	A	112	SER	3
1	A	46	LEU	3
1	A	63	VAL	3
1	A	94	SER	2
1	A	67	LYS	2
1	A	151	LEU	2
1	A	28	SER	2
1	A	76	ASP	2
1	A	100	VAL	2
1	A	16	VAL	2
1	A	169	MET	2
1	A	127	ASP	2
1	A	103	SER	2
1	A	92	MET	2
1	A	96	ASN	2
1	A	170	GLN	2
1	A	117	VAL	1
1	A	18	ASP	1
1	A	12	LEU	1
1	A	55	THR	1
1	A	57	VAL	1
1	A	113	ASN	1
1	A	77	THR	1
1	A	155	LEU	1
1	A	174	SER	1
1	A	52	SER	1
1	A	35	ILE	1
1	A	7	ILE	1
1	A	116	ASP	1
1	A	138	ASN	1

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.6 Ligand geometry ⓘ

There are no ligands in this entry.

6.7 Other polymers ⓘ

There are no such molecules in this entry.

6.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

7 Chemical shift validation

No chemical shift data were provided